

Trends in Computer System Design

Jerry Huck
HP Fellow

© 2005 Hewlett-Packard Development Company, L.P.
The information contained herein is subject to change without notice



Agenda

- HP Computer Strategy
- sx2000 system design
- Power and Cooling
- Summary



A global leader in technology and services for today's enterprise

- Technology
 - Blades, grid, utility computing, Integrity servers, Information Lifecycle and Business Management software
- Business and IT services
- Platform share of Oracle, SAP, Microsoft and Siebel
- 65,000 service and support professionals
- Public Sector, Financial Services, Manufacturing and Network Solutions Providers
- Unprecedented reach
 - Enables more than 100 million mobile phone subscribers
 - Handles two out of three credit-card transactions
 - Powers 106 of the world's 120 stock exchanges
 - Supports 95% of the world's securities transactions
 - Provides LAN-based carrier communications for 18 million people
 - Ships approximately 1 million printers per week

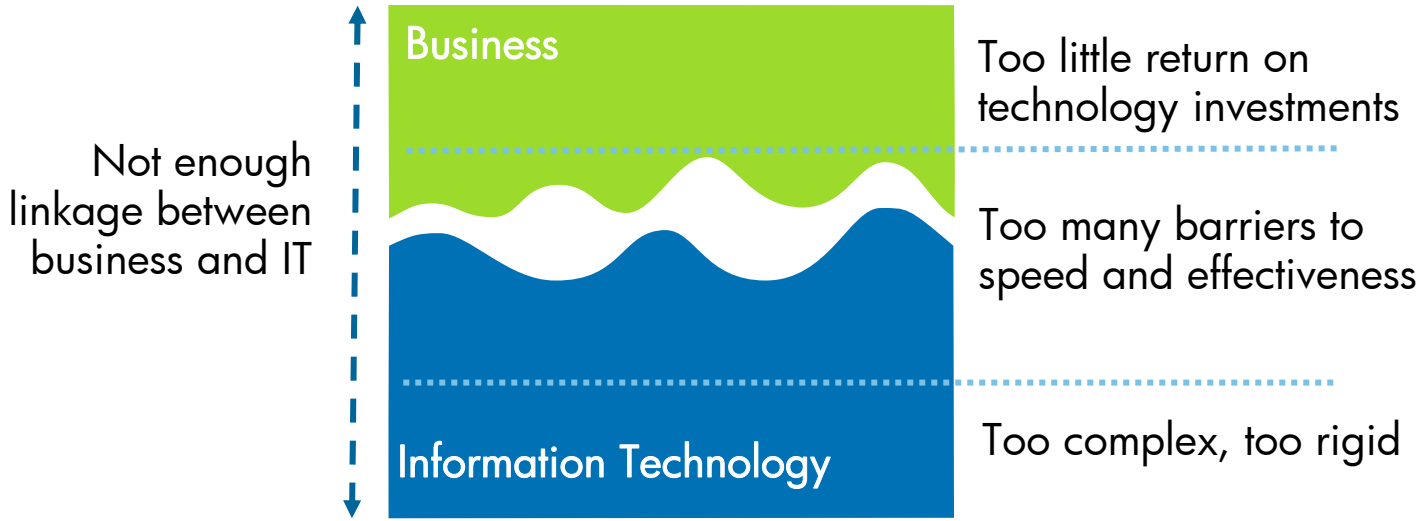






The Adaptive Enterprise

Is IT an enabler or inhibitor to change?



Business and IT synchronized to capitalize on change

Simplicity

- Reduce IT cost and complexity
- Make it easier to implement change
- Ensure resources work together

Agility

- Adapt in real time to business needs
- Drive change to create competitive advantage

Value

- Enable growth and profitability
- Unlock the value of assets
- Free up resources for innovation

Infrastructure for the Adaptive Enterprise

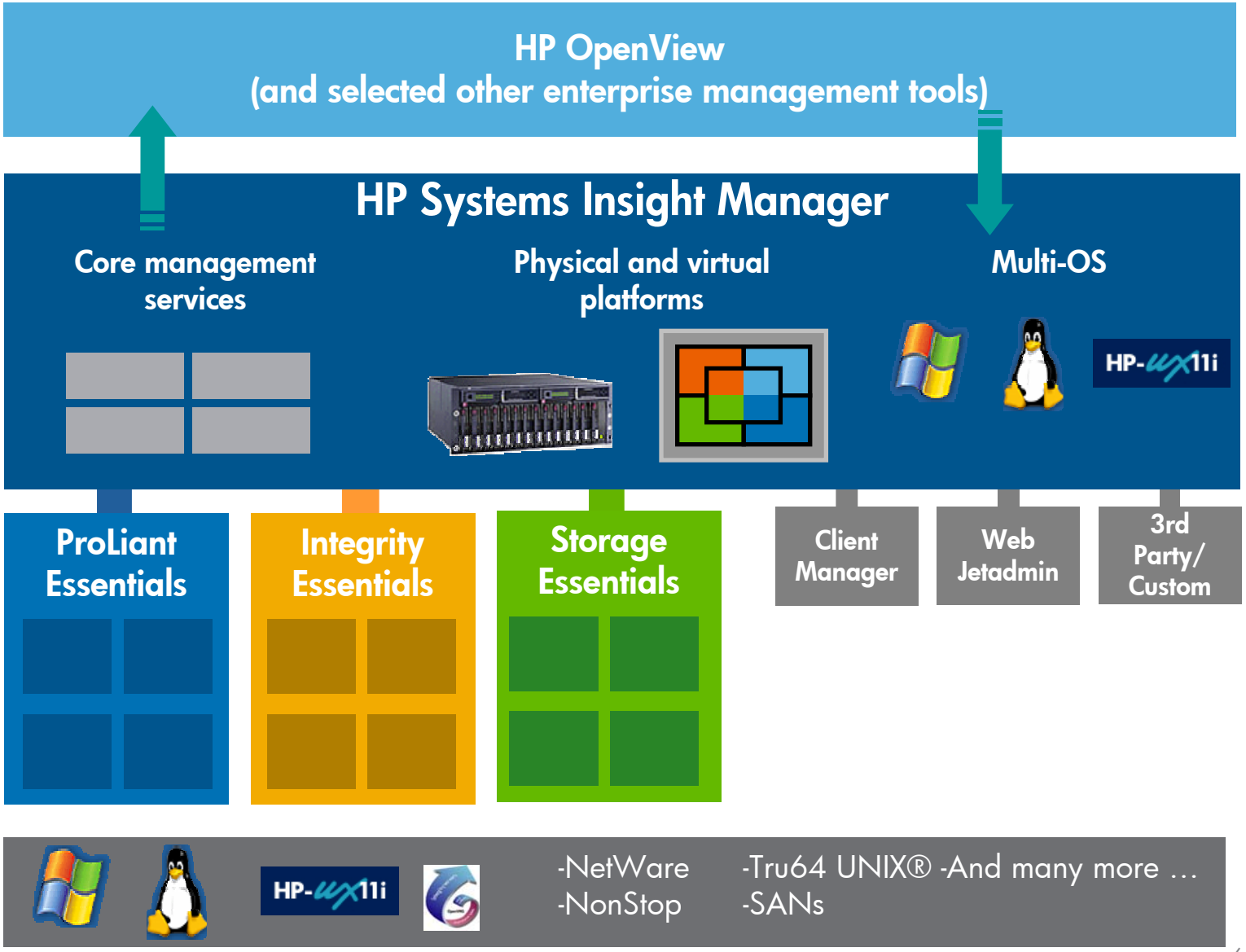
- **Modular**
- **Industry Standards based for low cost**
- **Added innovation to enable simplicity, agility, and value generation.**

- **System Design Challenge**



Unmatched breadth of solutions to your IT challenges

HP unified infrastructure management architecture



IT imperatives drive system design requirements for HP Integrity systems



Bring our high value projects up and online quicker



Better return on IT with HP Integrity systems

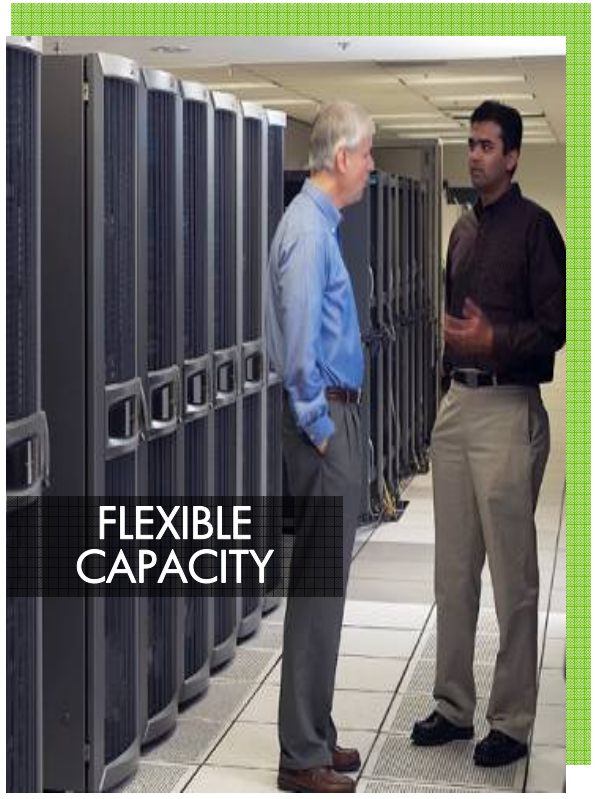
Consolidate and simplify to free up resources



Meet the high service level expectations of our customers

Flexible capacity for our changing business needs

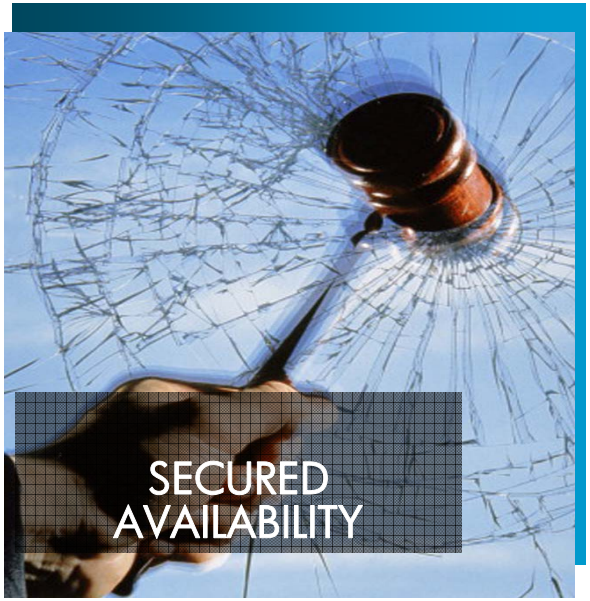
- Adjust to changes in real time
 - Move resources to where they are needed, easily and quickly
 - Handle workloads at peak times
 - Leverage spare CPU cycles
- Have the capacity for all workloads
 - Handle transactions, batch, and queries
 - Enough resources for ad-hoc workloads
 - Immediate response for real-time requests
- Scale systems as our needs change
 - Leverage modular building blocks
 - Repurpose systems as application needs evolve
 - Pay for contingency capacity if and when needed



HP Integrity servers deliver
real-world performance and flexibility for today's diverse workloads

Secured availability to preserve service levels

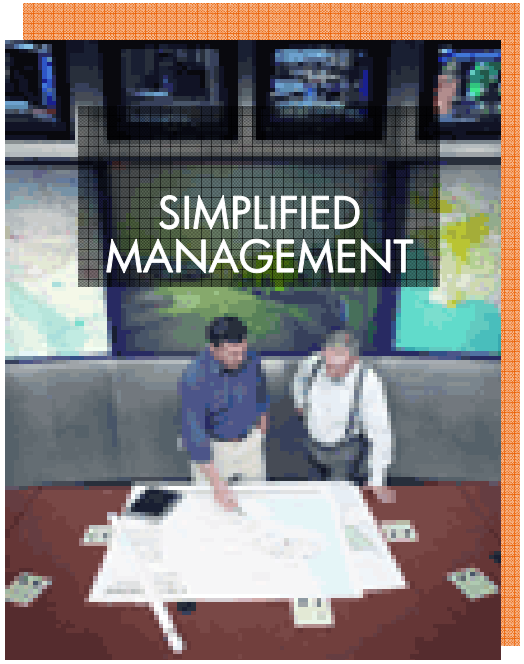
- Support continuous business operations
 - Maximize application availability to the users
 - Minimize planned and unplanned downtime
 - Recover quickly from a disaster
 - Preserve data integrity
 - Optimize user response time
- Protect our IT environment at every level from intruders
 - Keep intruders out while providing authorized users what they need
 - Secure the full environment, not just parts of it



Integrated availability and security solutions for HP Integrity servers keep your applications up and running, preventing business operations from being affected by IT outages…whatever the cause

Simplified management to streamline operations

- Improve productivity of system managers and DBAs
 - Make managing our IT as consistent as possible
 - Implement updates and patches faster and easier
 - Automate to save time and eliminate human error
- Optimize the use of our datacenter facilities
 - Allow us to add capacity without adding floor space
 - Keep our power and cooling costs under control
- Leverage industry best practices for efficient IT operations
 - Streamline our IT service operations
 - Make vs. buy decisions



Powerful, integrated management tools for HP Integrity servers simplify management of your IT operations so you can focus on your business

Major Trends and Challenges – context for the sx2000 design



- Multi-core Designs
 - Dual core common
 - Talk of more, sometimes much more
 - Sun delivering 8 cores in T1
- Chipsets provide path for hardware differentiation
- Power density going up

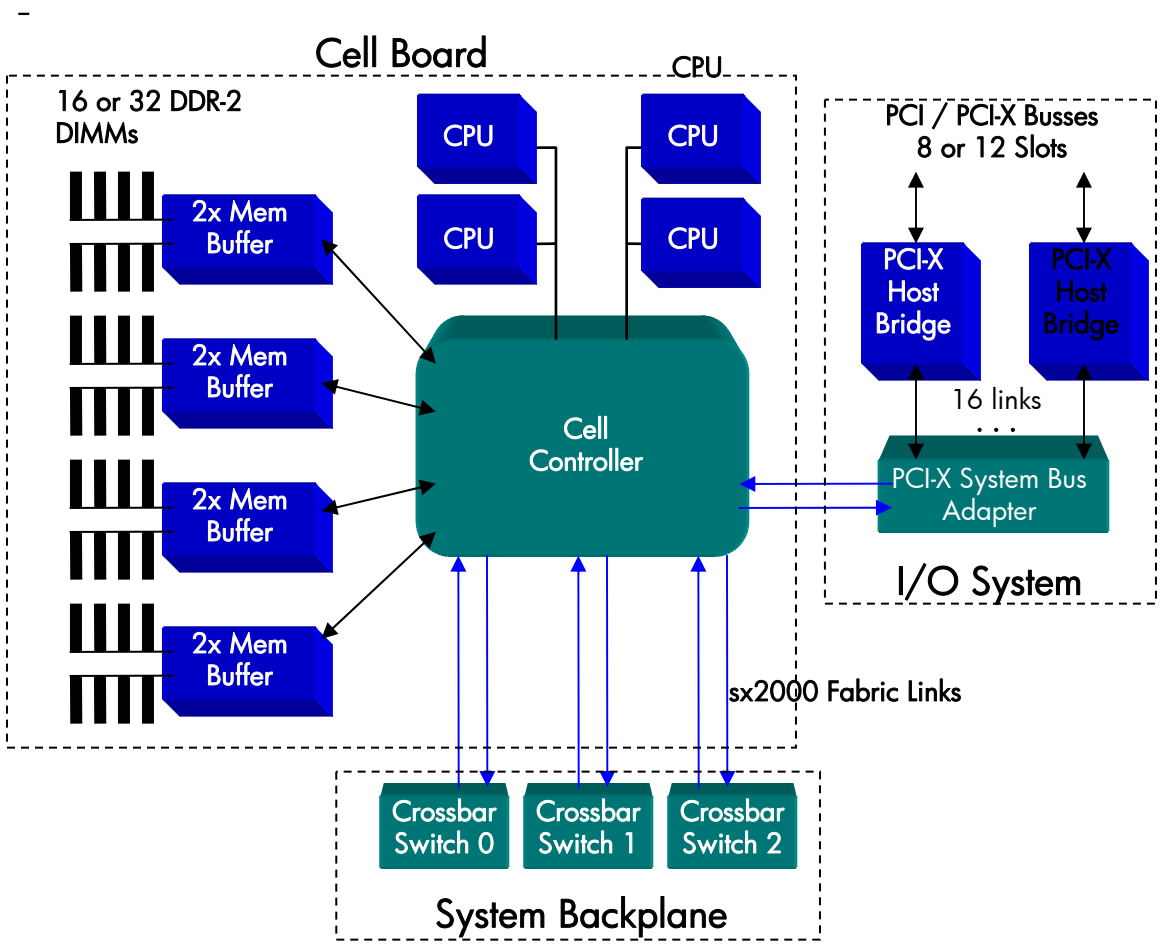
HP Integrity servers

New sx2000 chipset

64	HP Integrity Superdome	 Up to 64 processor scalability and hard-partitioning capability for leading consolidation • Up to 64 Intel® Itanium® 2 processors • Up to 1 TB memory • 192 PCI-X slots (with I/O expansion) • Up to 16 hard partitions
16	HP Integrity rx8640 server with server expansion unit (SEU)	 16 processor scalability and hard-partitioning capability for consolidation • 2 to 16 Intel Itanium 2 processors • Up to 128 GB memory • 32 PCI-X slots (with SEU) • Up to 4 hard partitions
8	HP Integrity rx7640 server	 8 processor flexibility with high-performance, density, and hard-partitioning capabilities • 2 to 8 Intel Itanium 2 processors • Up to 64 GB memory • 15 PCI-X slots • Up to 2 hard partitions
8/4	HP Integrity rx4640 server	 4/8 processor high-performance server in ultra-dense form factor • 1 to 4 Intel Itanium 2 processors or 2- to 8-way HP mx2 dual-processor modules • Up to 128 GB memory • 6 available PCI-X slots • 10 servers per 2m rack
	HP Integrity rx2620 server	 2 processor high-performance, server for multi-purpose entry-level computing • 1 to 2 Intel Itanium 2 1.3Ghz/3MB, 1.6GHz/3MB. 1.6Ghz/6MB processors • 2U form factor • Up to 24 GB memory • 4 PCI-X slots • 20 servers per 2m rack
2	HP Integrity rx1620 server	 2 processor ultra-dense server delivering maximum computing value • 1 to 2 Intel Itanium 2 1.3 GHz/3MB, 1.6GHz/3MB/533MHZ FSB processors • 1U (rx1620-2) form factor • Up to 16 GB memory • 2PCI-X slots • 40 servers per 2m rack
	HP Integrity Blades	New 1 or 2 processor Itanium2 – based blade • 1 to 2 Intel Itanium 2 1.6 GHz/3MB processors, • Up to 8 GB memory

What is a chipset?

The chipset plays a vital role in providing the interconnectivity between processors, DRAMs and I/O cards, turning this group of components into a high performance computer system.

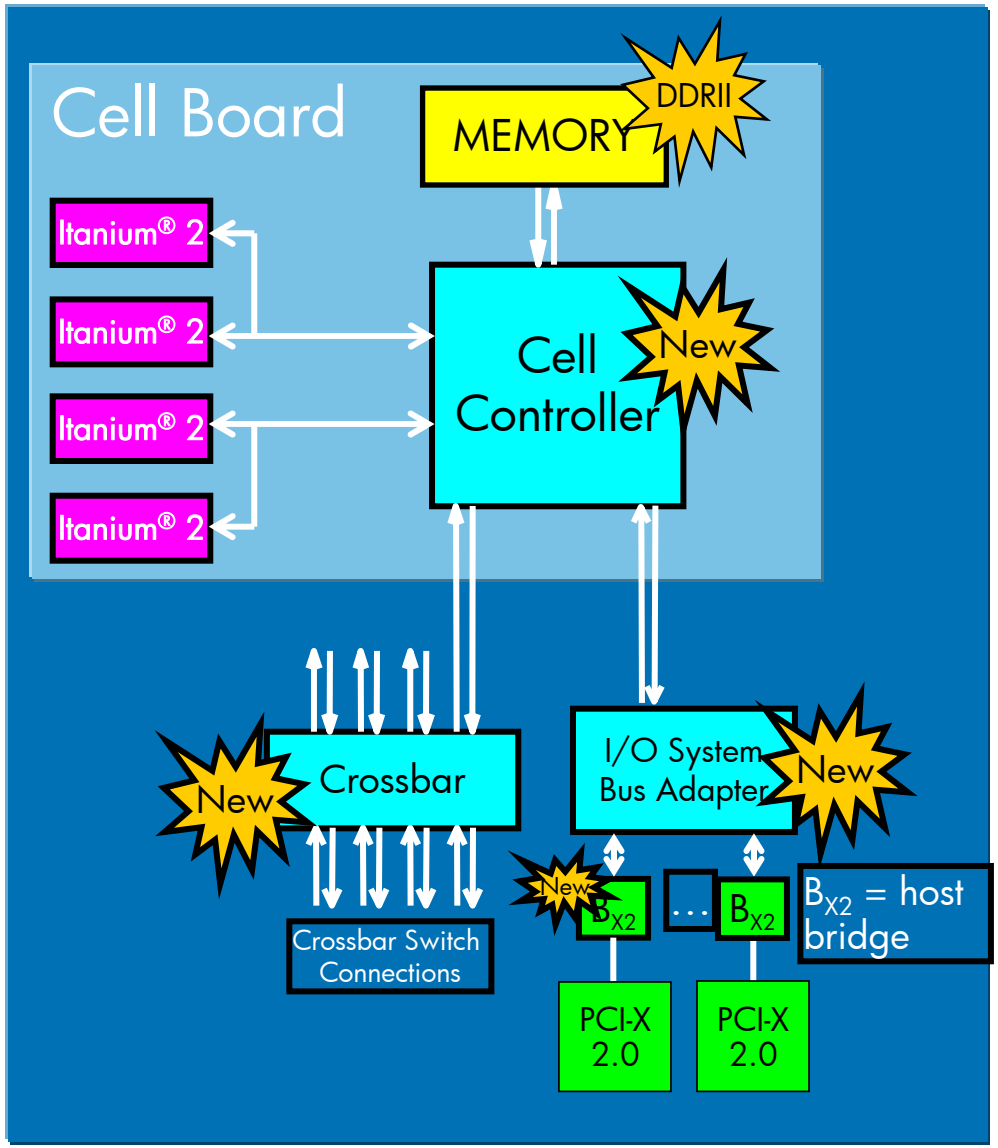


sx2000 Systems

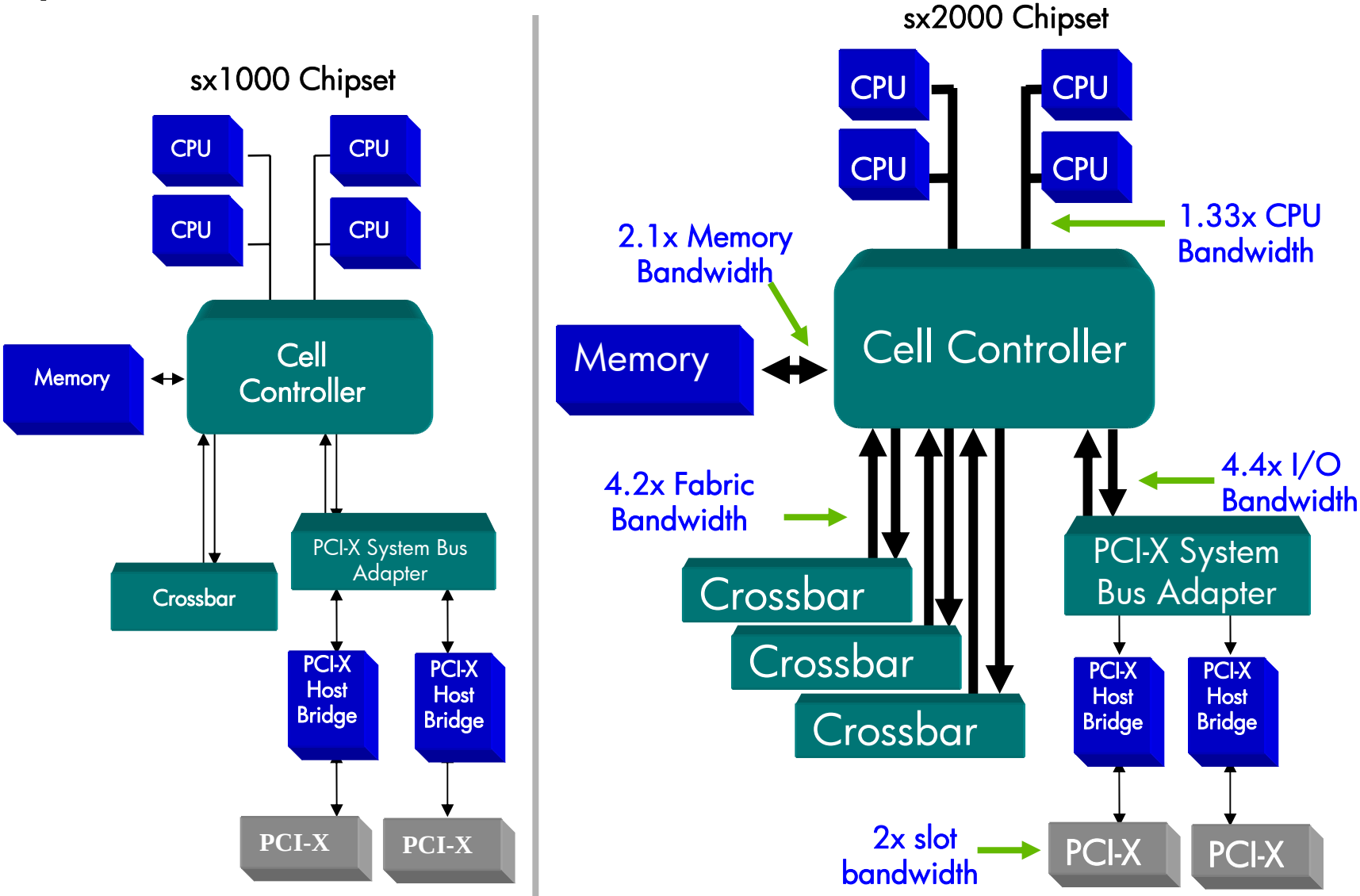
- Major refresh of HP's High-end and Midrange product families
 - Integrity Superdome, rx86xx/rp84xx and rx76xx/rp74xx systems
 - New shipments and in-box field upgrade for these systems
- Optimized around next generation processors
 - Itanium: Itanium 2 9M, "Montecito" and "Montvale" processors
 - PA-RISC: PA-8900 processors
- Powered by sx2000 chipset
 - Significant improvements in performance and reliability, availability, serviceability, and manageability (RASM)
- Multi-OS support:
 - HP-UX 11i v1/v2/v3, *Windows Server 2003 SP1 and Longhorn*, Red Hat Linux Advanced Server 3.0 (Montecito and Montvale), SUSE Linux (Montecito and Montvale), OpenVMS 8.3

What's new in sx2000 systems

- Memory
 - DDR2 memory technology
 - Increased memory address space
 - Enhanced DRAM ECC
- I/O
 - PCI-X Mode 2
 - PCI Express in future release
- System Interconnect
 - New link technology for fabric and I/O
 - Upgraded crossbar topology
- Performance improvements
 - Increased bandwidths, reduced latencies
- I/O Write coalescing
- Fail-over system clock for high reliability



The sx2000 provides greater performance with increased bandwidth



sx2000 provides significant bandwidth improvements

	sx2000 chipset peak (sustained)	Improvement
Processor buses / cell	17.0 (13.6) GB/sec	1.33x
Memory subsystem / cell	17.2 (16.0) GB/sec	2.1x
Fabric link / cell	34.6 (27.3) GB/sec	4.2x
I/O link / cell	11.5 (8.0) GB/sec	4.4x
I/O single slot	2.0 GB/sec	2X

Reduced system latencies with the sx2000 chipset

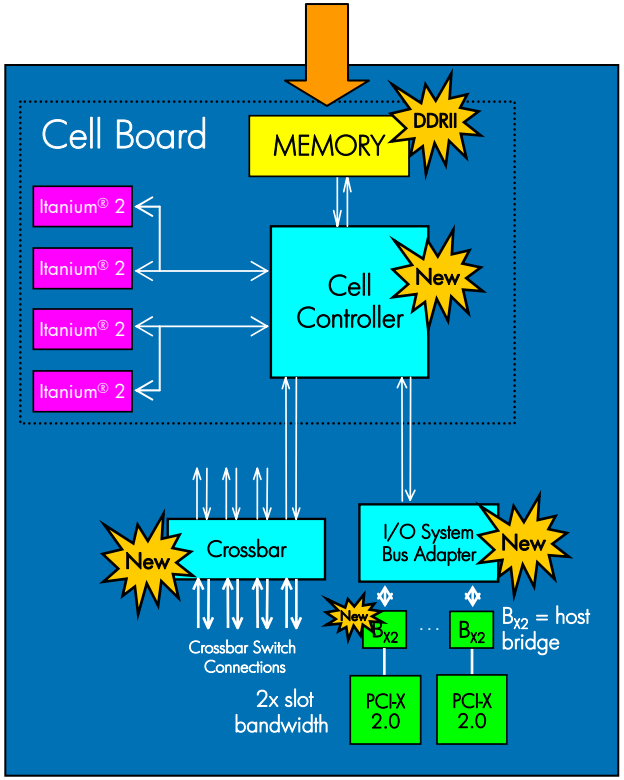
	Reduction
Memory Latency (load-to-use):	
4-core interleaved or Cell local	-23%
8-core interleaved	-37%
16-core interleaved	-33%
32-core interleaved	-20%
64-core interleaved	-10%
Cache-to-Cache Latency (load-to-use):	
4-core	-30%
8-core	-44%
16-core	-45%
32-core	-32%
64-core	-18%

Greater performance and availability with memory advancements

Memory advancements with the sx2000:

- Capacity: 4x improvement
 (DDR2 SDRAMs with support for 2Gb SDRAMs when available)
- Bandwidth: 2x
- Latency: > 25% reduction in memory access latency
- Availability: Double chip-sparring
 (Memory error correcting logic to re-establish chip-spare correction after another DRAM in the same ECC codeword has failed.

 Eliminates system downtime to replace failed DRAMs)



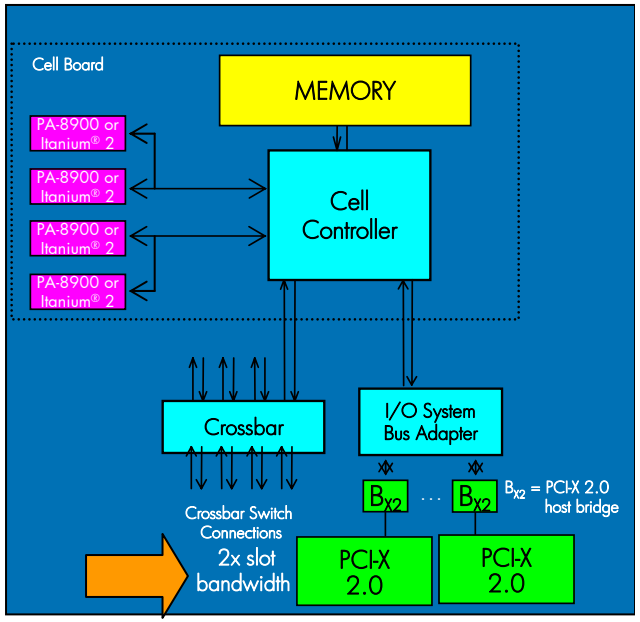
Enhanced I/O connectivity

I/O advancements with the sx2000:

- Cache Size: 50% increase improvement (ensures increased bandwidth of I/O cards can be satisfied)
- Bandwidth: 4x expansion
- Reliability: Enable ECC (Error correcting code) for any PCI-X card that supports ECC
- Cards: Support for leading edge 10GB Ethernet cards, Fiber channel cards, Infiniband cards, SCSI, Gigabit Ethernet, Fiber Channel, multi-port/combination, and more

 Support for legacy I/O cards

PCI-Express: Planned for 2006



Greater availability and reliability with the high-speed links re-design

- **Greater Reliability**: Tolerating transient errors
 (When an error is detected on a high-speed link, any transactions in flight will be re-transmitted with the correct data.)

- **Greater Availability:** Tolerating a hard error in one of the channels that make up the link.
 (If the chip determines that one of the channels of the link has gone "bad", it will swap in the spare channel and continue running at full bandwidth.)

- If a link or crossbar chip fails completely, the sx2000 **minimizes downtime** by allowing the partitions affected to be immediately rebooted using the unaffected fabrics.

sx2000 system reliability and availability features

Location	Feature
Memory system	Error detection / correction DIMM address parity protection Double chip sparing
Processor	Cache error detection / correction
Processor bus	Error detection / correction
I/O and fabric links	Link level retry Link spare channel
Crossbar / System fabric	Redundant links to cells
I/O slots	Error detection / correction PCI failure isolation to a single slot
Chipset	Internal data path error detection / correction
Partition	nPars – Hardware isolation between partitions GSM communication without loss of isolation Cell OLARD PCI card OLARD

Greater performance, availability and investment protection with Integrity sx2000-based systems

Performance & scalability

- More powerful processors
- 33% CPU bus bandwidth increase
- 2x memory bandwidth increase
- 4x memory capacity
- 4x I/O bandwidth increase
- 2x I/O slot bandwidth increase
- 4x crossbar bandwidth increase
- 25-33% latency reductions







Investment protection

- Field upgradeable into existing customer systems
- Fits within the available power / thermal windows of existing systems

Increased single system high availability

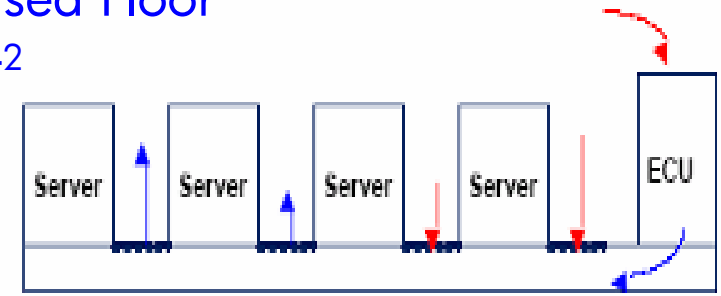
- DRAM Erasure: memory is System Bus Error (SBE) protected even after chip fail
- Multi-path crossbar topology
- Link level retry with 'self-healing' on crossbar and I/O links
- Greater fault protection within the CPU (cache repair-on-the-fly)
- Redundant DC/DC converters and system clocks
- Less than 15 minute removal and replacement of a FRU for one person after failed FRU is identified and before boot sequence begins

Power and Cooling

Problem: What HP found in 1998

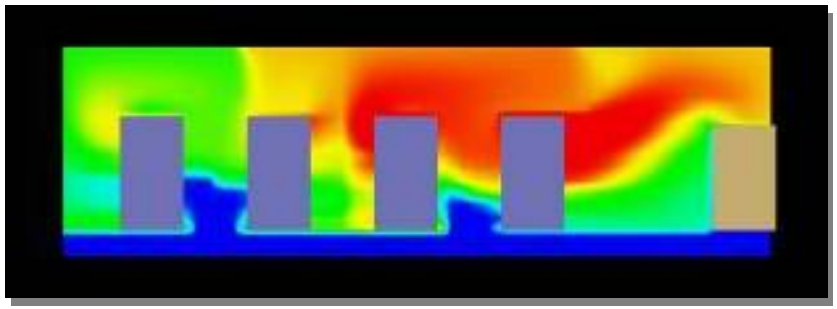
Data Center Density

- Existing Data Centers: 40-70 W/ft² & 18" Raised Floor
- Recommend: 100 W/ft² & 36" Raised Floor
- Server Roadmap: 75-225 W/ft²



Airflow Issues

- Distribution
- Inadequate air flow from tiles
 - Nominally 200 - 700 cfm
 - Servers require much more
- Cable congestion



An industry-wide problem!

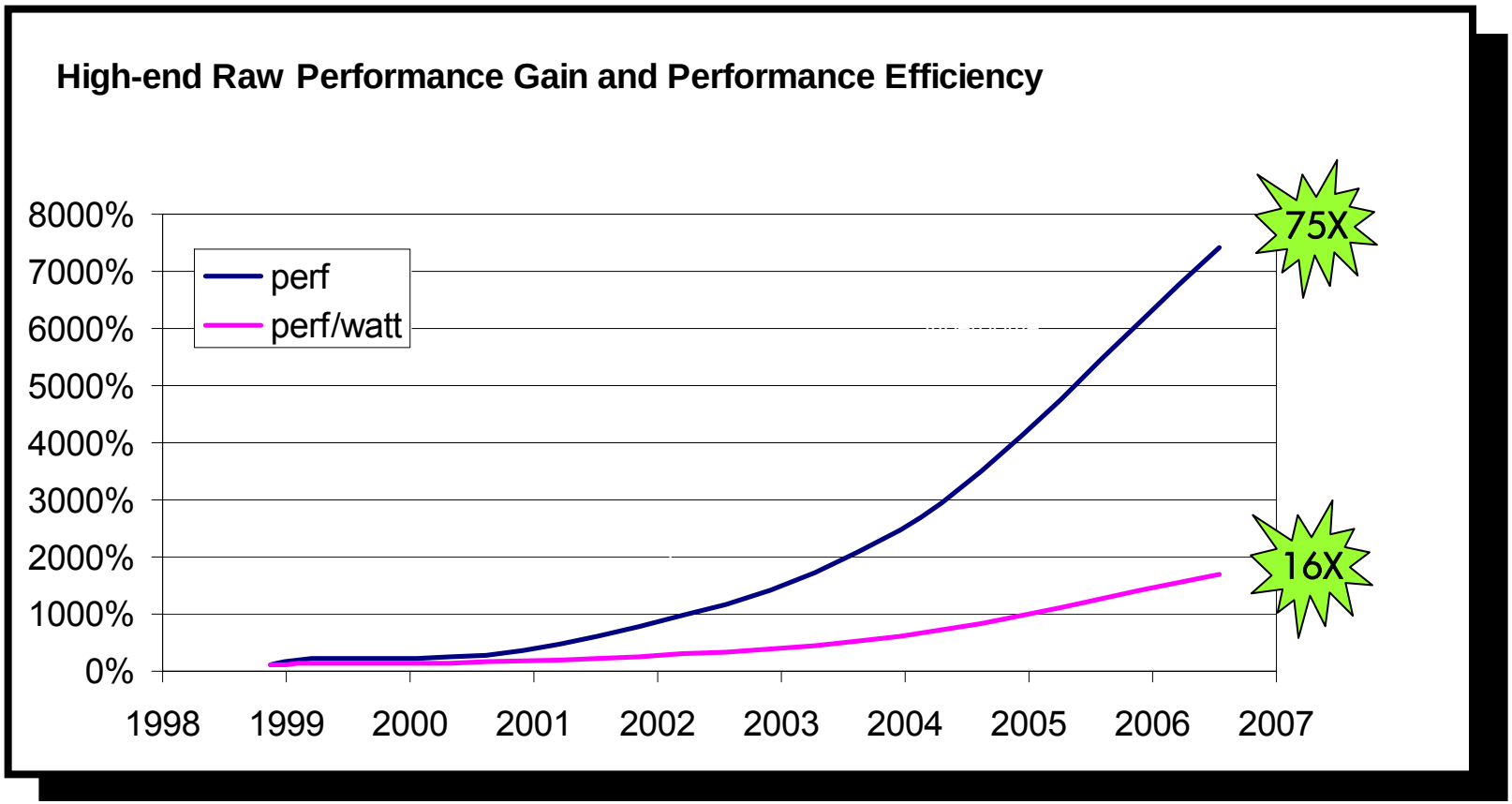
Example: Tip of the Iceberg

Source: Belady/Malone, HP

HP's Focus On Performance Efficiency continues to build

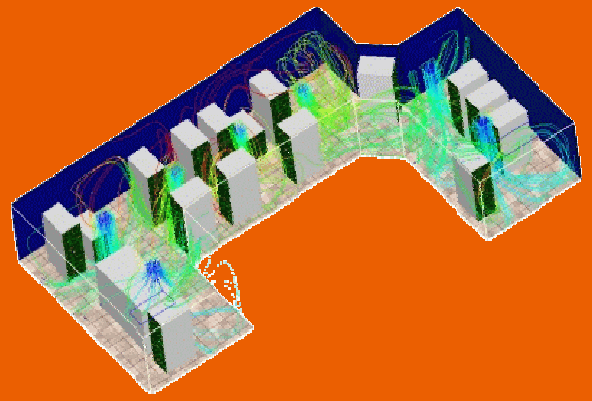
Approximate increases by 2006 platforms since 1998

- 75X performance increase
- 16X performance efficiency



Source: Belady/Malone, HP

Data Center Thermal Design



- Best Practices
- Commercially available Tools
- HP's Static and Dynamic Smart Cooling

invent

Following Best Practices will result in:

- Lower overall server temperatures and thus increase the reliability of your servers
- Extend the useable life of your data center
- Maximize the server density in your data center
- Significantly lower the energy costs in your data center
- Give you more power to use for computations
- Maximizes your uptime!

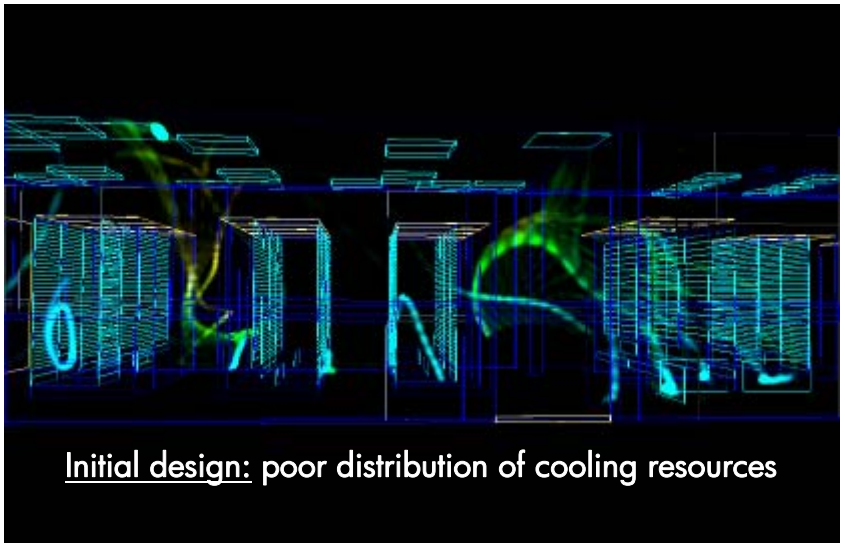
Source: Belady/Malone, HP

Some Examples of Best Practices:

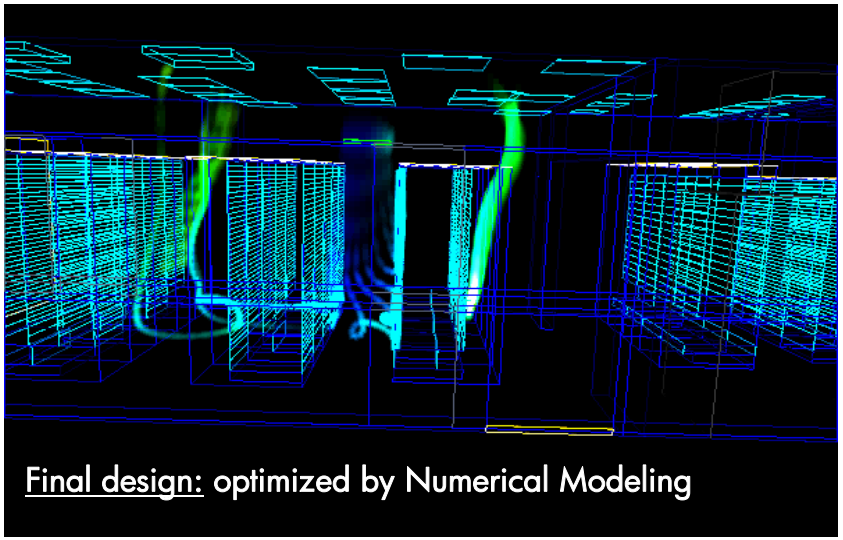
1. Hot Aisle/Cold Aisle
2. Matching Server Airflows
3. Eliminate Gaps in Rows & use longer rows
4. Use Cabinet Blanking Panels
5. Orient AC units perpendicular to hot aisles
6. Seal cable cutouts
7. Use 0.8 to 1.0 Meter High Floors
8. Use High Density and Low Density Areas
9. Use CFD Modeling for final optimization

Source: Belady/Malone, HP

Static and Dynamic Smart Cooling



Initial design: poor distribution of cooling resources



Final design: optimized by Numerical Modeling

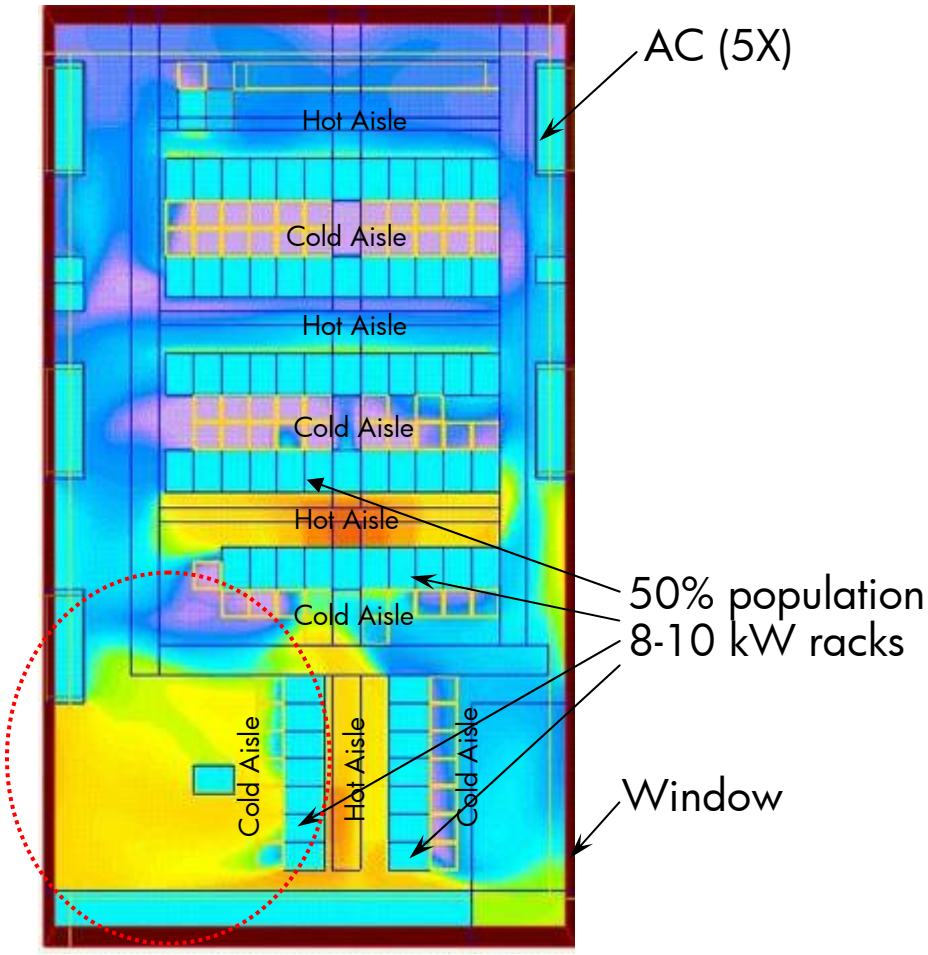
- Matching of AC resources to IT equipment requirements by modeling, metrology and intelligent control
- Two key components:
 - *Static Smart Cooling**
 - *Dynamic Smart Cooling*
- Target cooling resources energy savings (compared to state of art)
 - 25% with static provisioning
 - 50% with dynamic
- Target payback ~ 1- 2 years

** Available world-wide*

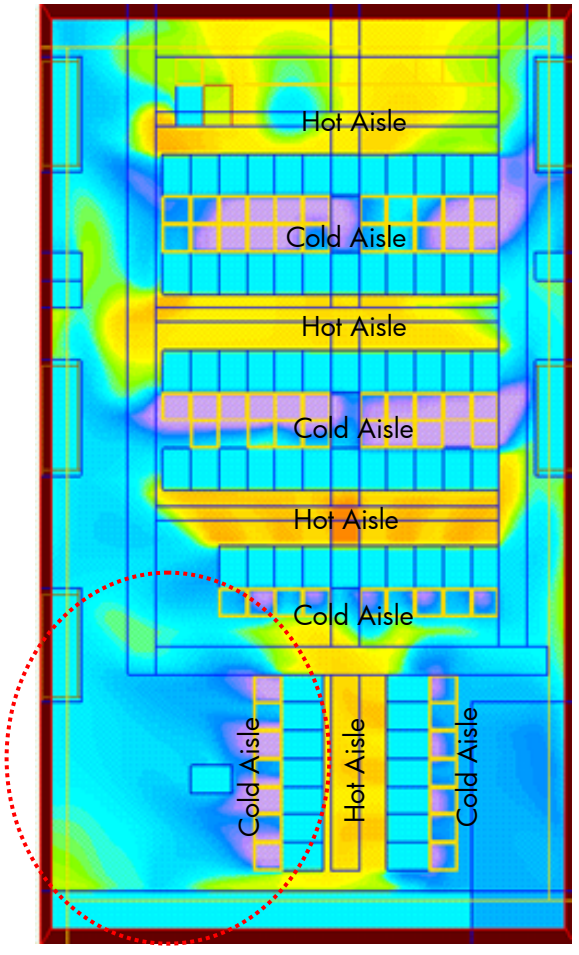
Source: Chandrakant Patel, HP

Static Smart Cooling
Customer Example

Initial Design:
- Uniform 150 W/ft^2 cooling
- Hot spots from dense distribution of 1U racks



HP Solution:
- CFD modeling to strategic repositioning inlet air vents - the only degree of freedom
- Equipment upgrades also considered



Source: Chandrakant Patel, HP

Novel Data Center Cooling Technologies



HP-Developed Dense Data Center Cooling Solutions – First in the Industry!

In 2000 HP built a Prototype DC to test Data Center cooling capacity in a controlled environment.

- Measured both raised floor & DataCool capacity
- Measured transient response of data center
- Consultants called this work unprecedented



Co-developed DataCool: high density data center cooling technology

- With Ericsson Power (sold to Liebert)
- First new DC Technology in years
- Tested
 - Raised floor capacity: 175 W/ft²
 - DataCool capacity: 475 W/ft²
 - Combined capacity: 560 W/ft²

Water Based
single phase
20kW
8 sliding fans
6ft x 6ft



Product Evolution

Latest generation: XDO

- Owned and sold by Liebert
- Highest density air cooling system
- Combined capacity: ~900 W/ft²

R134a refrigerant
2 phase
32kW
3 fixed fans
6ft x 6ft



Source: Belady/Malone, HP

Other Dense Data Center Cooling Solutions

- Since DataCool, many other solutions have emerged.
 - Most are liquid cooled rack level available from IBM, SGI, Fujitsu, Rittal…
 - Some of those include:

Fujitsu Siemens 20kW Rack



IBM Rear Door Heat eXchanger 15 kW Rack



SGI Altix 3700 Liquid Rack (64 Itaniums/cabinet)



Source: Belady/Malone, HP

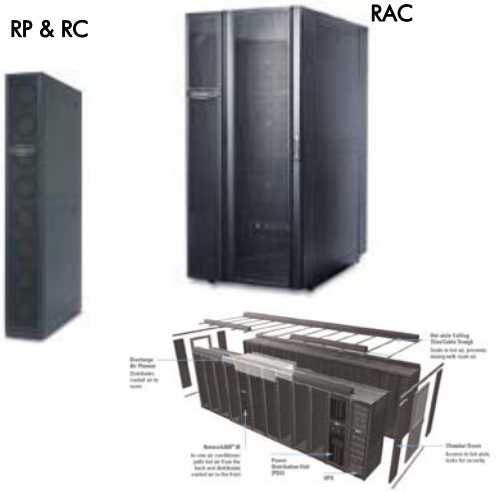
Other Dense Data Center Cooling Solutions

- Room Cooling Suppliers have emerged with new solutions and options as well…
 - Some of those include:

HP Modular Cooling System (MCS)



APC High Density Products

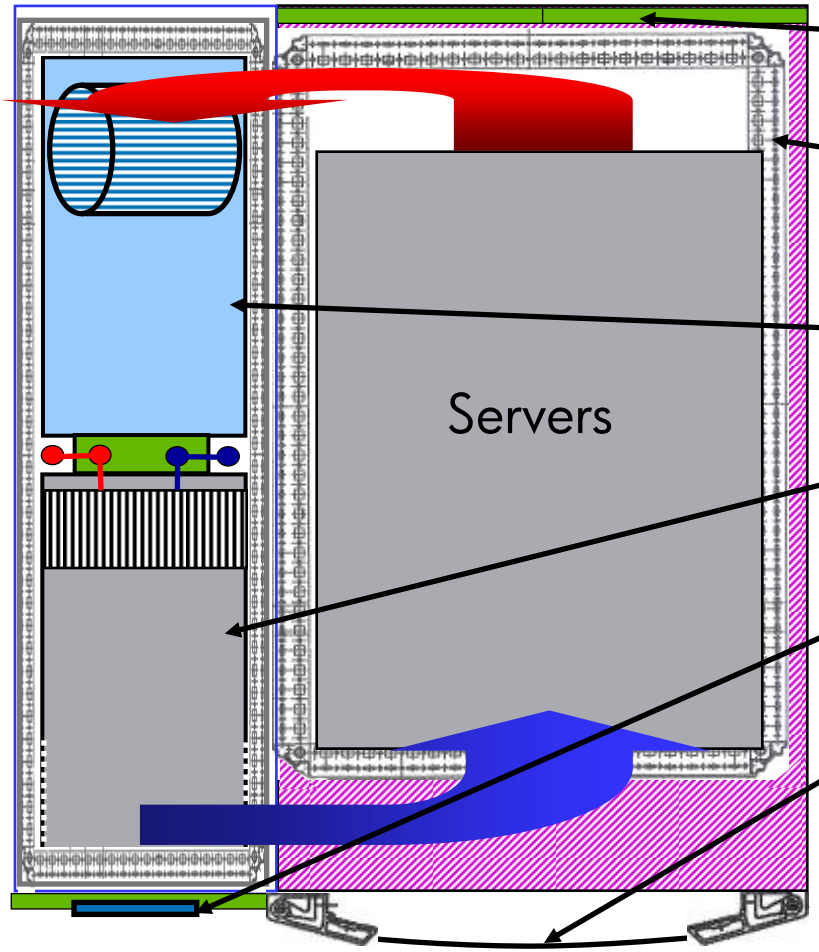


Liebert High Density Products



Source: Belady/Malone, HP

MCS Design Layout

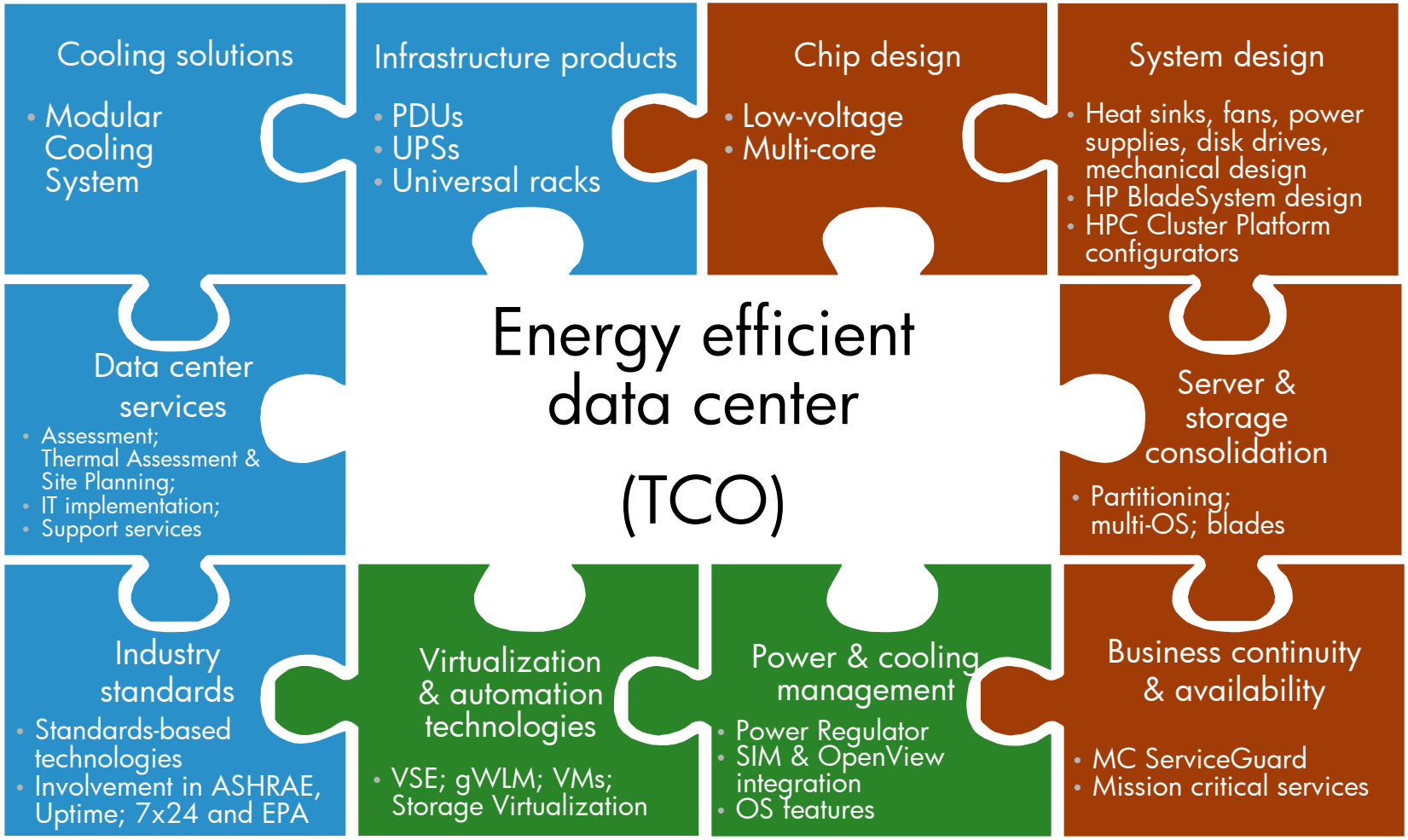


Top View

- Sealed Split Rear Door
- Standard 10642-G2 rack
- Fan Assembly FRU
- Heat Exchanger FRU
- Front Panel Display
- Sealed Front Door with add-on plenum
- Can be adapted to either side of the rack

Source: Belady/Malone, HP

Complete portfolio to optimize data center efficiency



Putting the portfolio to use

European banking customer

Business needs	HP solution	Customer benefits
• 5000 blades already deployed • Running out of power and cooling capacity • Space constrained • Performance hungry	• DC assessment & planning services • Revamped data center to best practices • Future MCS implementation in new data center • Integrated power and cooling planning	• Installed a further 1000 blades • Improved server density by 31% • Saved $500k on infrastructure changes • New data center plans to deploy 10,000 blades in a 22kW rack, cooled by HP's MCS
Inflexible power and cooling infrastructure	Integrated power and cooling planning	Improved data center efficiency (TCO)

Summary

…More capabilities – announced 20 March 2006



- More flexibility: VSE Acceleration
- More flexibility: Global Instant Capacity
- More capacity: Faster sx2000 chipset
- More capacity: rx7640, rx8640, Superdome



- More availability: RAS features in sx2000 chipset
- More availability: Low cost, HP Integrity NonStop NS14000
- More availability: Better disaster tolerance for HP-UX 11i
- More availability: Leading multi-OS SAP clustering

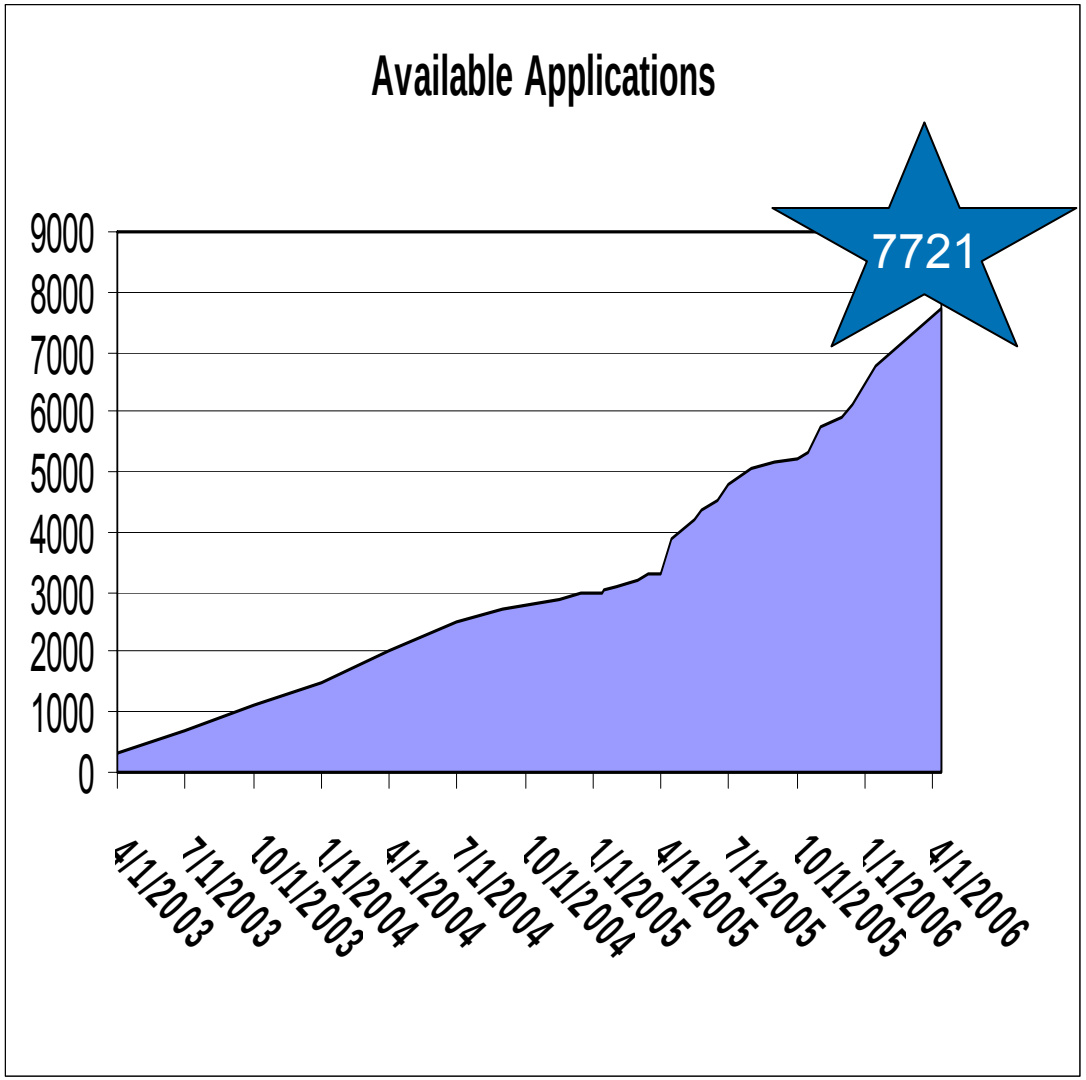


- More simplicity: More efficient HP-UX 11i management
- More simplicity: Application auto-discover for VSE

HP Integrity application support – our #1 Priority Making Strong and Steady Progress

HP Integrity servers at:

- One of the largest manufacturing installations of SAP in the world
- One of the largest instances of PeopleSoft in the world
- Largest wireless telco provider in North America
- One of the largest database installations of Microsoft SQL Server in North America
- One of the largest Windows- based Siebel Public Sector deployments in Asia Pacific



HP Integrity server: Delivering value beyond the chip

Flexible capacity | Secured availability | Simplified management

Horizontal solutions
- Business Intelligence
- Infrastructure
- High Performance Computing

Applications

Middleware

Operating system

Hardware

Intelligent control

Assess, Advise, Automate

Virtual servers

HP Virtual Server Environment

Vertical solutions
- Financial services
- Manufacturing
- Network & service providers
- Public sector

Operating systems

HP-UX 11i

OpenVMS

Windows Server 2003



HP Integrity servers





invent